Journal of Computer Science

Original Research Paper

A New Hybrid Model to Predict the Performance of Trainee Teachers Based on Clustering and Classification

[1]Yissam Lakhdar, [2]Khawla EL-Bendadi and [1,3]Brahim Bakkas

[1]*Department of Computer Science, Regional Center for Education and Training Professions, Fez-Meknes, Morocco*
[2]*Department of Mathematics and Computer Science, École Marocaine des Sciences de l'Ingénieur, EMSI, Fez, Morocco*
[3]*Department of Computer Science, École Nationale Supérieure d'Arts et Métiers (ENSAM-Meknès)*
Moulay Ismail University, Meknes, Morocco

Article history
Received: 20-04-2024
Revised: 13-05-2024
Accepted: 29-05-2024

Corresponding Author:
Yissam Lakhdar
Department of Computer Science, Regional Center for Education and Training Professions, Fez-Meknes, Morocco
Email: lakhdar_yissam@yahoo.fr

Abstract: This article explores how artificial intelligence, particularly machine learning, can be used to assess and predict the performance of trainee teachers in Morocco. Considering the country's focus on integrating technology into education and the challenges posed by the COVID-19 pandemic, the authors propose a novel hybrid model that combines clustering and classification algorithms. This model aims to understand the Information and Communication Technologies (ICT) skills of trainees from various backgrounds and predict their performance after training at a Moroccan center. It should be noted that, in this study, we used trainee data in compliance with ethical principles and confidentiality protocols. All data collected were anonymized in order to protect the identity of the participants and guarantee their confidentiality. The study investigates whether a gap exists in the digital literacy of trainees based on their prior degrees and analyzes their progress after training. By applying the hybrid model, the research identified distinct groups of trainees, including high achievers and a mixed group with varied performance. The findings suggest that while a trainee's digital skills may be influenced by their prior institution, the training program effectively improves their ICT skills and allows them to achieve success. The clustering algorithm used prior to classification provides a better understanding of the data and improves the classification rate. The experimental results provide valuable information for scientists looking to take advantage of new clustering techniques and classification for a variety of applications in data analysis. The paper further explores the impact of AI in education, details the proposed model, and discusses the results alongside potential avenues for future research.

Keywords: Teaching, Academic Performance, Learner, Machine Learning, Evidential Approach

Introduction

Education is a concept that highlights the acquisition of knowledge and values, participating not only in the development of people but also in the country's progress (Hoot *et al*., 2004; Lamichhane, 2018). Given the importance of this area, educational organizations always aim to provide a high-quality education that preserves school dropouts and the delay of learners in their studies.

For several years, the Ministry of National Education (MNE) in Morocco has focused on the integration of Information and Communication Technologies (ICT) in the teaching-learning process by equipping all schools with multimedia equipment and linking them to the internet and by setting up training programs within the framework of a program named generalization (GENIE) of Information and Communication Technologies in Education (ICTE) in Morocco based on two modules: Introduction to computer science and pedagogical use of ICTE (Raouf *et al*., 2020). This strategy has brought changes to the initial teacher training programs in the training centers by adopting a module integrating ICTE (Nejjari and Bakkali, 2017). As part of the Moroccan development model introduced in 2021, the education system is undergoing expansion and transformation in alignment with the strategic vision of the 2015-2030

© 2024 Yissam Lakhdar, Khawla EL-Bendadi and Brahim Bakkas. This open-access article is distributed under a Creative Commons Attribution (CC-BY) 4.0 license.

reforms (HCETS, 2015). These reforms were initially unveiled on May 20, 2015, by the Higher Council of Education, training and Scientific Research, along with the enactment of framework law 51.17, notably encompassing articles 33 and 42 (Berrada *et al*., 2022). The overarching goal for Morocco is to enhance the quality of education by integrating Information and Communication Technology (ICT) into teaching and learning processes, with a focus on placing learners at the forefront of education through training initiatives. Efforts are underway to modernize educational content and methodologies, drawing on the expertise of professionals in the field. Notably, the education sector has faced considerable challenges due to the impact of the COVID-19 pandemic, prompting significant shifts in teaching methods and tools. This crisis has underscored the imperative for both teachers and learners to acquire digital literacy skills (Bachiri and Sahli, 2020; Mounjid *et al*., 2021; Outoukarte *et al*., 2023). Consequently, stakeholders in the education sector have identified various challenges associated with the use of ICT in education, sparking numerous questions and concerns.

As previously mentioned, the Ministry of Education emphasizes the integration of ICT into teacher training. These data-driven initiatives prioritize student-centered learning through regular assessments that identify students' strengths and areas for improvement, allowing teachers to tailor their approach to individual needs. Assessing learners' progress holds significant importance, prompting the implementation of diagnostic assessments at the onset of each academic level. These assessments enable teachers to discern the accomplishments and challenges of their students, facilitating the customization of teaching approaches to cater to their unique and individual requirements. Artificial intelligence, particularly machine learning, offers groundbreaking tools for learner assessment in schools. It can fundamentally reshape education by transforming how we teach, learn, and conduct research. Educational institutions like schools and colleges should seize this opportunity to leverage machine learning's potential. By effectively identifying struggling learners, they can personalize support and develop targeted action plans for improved success (Albreiki *et al*., 2021). There are various machine learning methods, including supervised learning, unsupervised learning, deep learning, and reinforcement learning (Nafea, 2018; Zhou, 2021).

The objective of this article is to predict the performance of trainee teachers by employing a novel hybrid model. The hybrid models, which combine clustering and classification algorithms, have appeared as a promising alternative method for dealing with the complex, multi-faceted nature of educational data. This model involves the initial application of an unsupervised learning algorithm, followed by the validation of the model using six supervised learning algorithms. The data utilized for this study were obtained from a dataset derived from two assessments digital Diagnostic Assessment (DA) and a Summative Assessment (SA). These assessments were administered to 135 trainee teachers at the Regional Center for Education and Training Professions (RCETP) in Fez-Meknes, Sefrou, Morocco.

The aim of this study is to develop a general understanding of the ICT prerequisites of primary school teacher trainees and to assess trainees' performance after 34 h of training at the RCETP, Fez-Meknes, Sefrou. The main objective of this study, by applying the hybrid model, is to predict trainees' academic performance and skill levels.

We hypothesized that:

- Hypothesis 1: Trainees' ICT skills depend on their training at their home institution
- Hypothesis 2: Trainees' ICT skills are influenced using traditional methods adopted in the teaching-learning process
  during the three years of post-secondary education
- Hypothesis 3: Trainee teachers used digital platforms and resources during the COVID-19 pandemic

Materials and Methods

Education is an important factor that can positively change the situation of a country by adjusting to the new society's demands, new profiles of learners and teachers, new methods of learning and teaching, and finally the communication and digital revolution.

For this purpose, artificial intelligence is used in education to facilitate certain tasks like learning assessment, improving the feedback quality, creating personalized learning paths, preventing school dropout, and solving more or less complex problems formulated by learners. To develop innovative solutions based on AI in education, it is very important to have significant data and powerful algorithms allowing their analysis.

Data Clustering

Data clustering is a process of extracting patterns from large data sets, it is commonly an unsupervised method that organizes data into homogeneous groups based on similarity between patterns. Although classical clustering algorithms such as C-means only provide hard partitions, many variants are proposed by introducing fuzzy sets (Bezdek *et al*., 1984), possibility theory, and evidence theory based on the credal partition that allows measuring accurately the uncertainty of the assignment of an object to a cluster (Krishnapuram and Keller, 1996; Masson and Doneux, 2008). Evidential clustering allows objects with different characteristics to belong not only to a singleton cluster but also to a set of clusters called meta-cluster. The use of evidence-based methods improves the decision-making process after a learner assessment. The effectiveness of these algorithms depends on the data type handled and the expected objective. The evidential C-means method takes advantage of the Fuzzy C-Means (FCM) algorithm and Noise Clustering (NC) to describe the uncertainty between clusters (Bezdek *et al*., 1984; Chen *et al*., 2024; Dave, 1991; Sen and Dave, 2002). It allows data to belong to a singleton cluster, a noise cluster ∅, or a meta-cluster depending on its belief mass. ECM is based on the minimization of the following objective function:

$$J_{ecm}(M,V) = \sum_{i=1}^{n} \sum_{A_j \neq \emptyset} c_j^{\alpha} m_{ij}^{\beta} d_{ij}^2 + \sum_{i=1}^{n} \delta^2 m_{i\emptyset}^{\beta} \quad (1)$$

Under the following constraint:

$$\sum_{i=1}^{n} m_i(A_j) + m_i(\emptyset) = 1 \quad (2)$$

M denotes the matrix of belief masses and V represents the matrix of specific and meta-cluster centers α is a parameter used to control the degree of penalization for high cardinality subsets, b is a weighting exponent and δ is an adjustable threshold to detect noise. c_j^{α} is the dimensional center of cluster j, and m_{ij}^{β} is the belief mass of x_i at the cluster j.

Throughout the iterative process, the objective function J is minimized by Lagrange multipliers to provide the credal partition matrix M for objects and the cluster center matrix. The values of the belief masses m_{ij}^{β} are updated according to the following equation:

$$m_{ij} = \frac{c_j^{\frac{-\alpha}{(\beta-1)}} d_{ij}^{\frac{-2}{(\beta-1)}}}{\sum_{A_k \neq \emptyset} c_k^{\frac{-\alpha}{(\beta-1)}} d_{ik}^{\frac{-2}{(\beta-1)}} + \delta^{\frac{-2}{(\beta-1)}}} \quad (3)$$

And the cluster centers are given by the matrix V that is, the solution of the following linear system:

$$H_t V_t = B_t \quad (4)$$

where,

$$B = \sum_{i=1}^{n} X_{iq} \sum_{A_j \neq \emptyset} c_j^{\alpha-1} m_{ij}^{\beta}, l = 1, c\ q = 1, p \quad (5)$$

And:

$$H = \sum_i \sum_{A_{j \supseteq (\omega_k, \omega_l)}} c_j^{\alpha-2} m_{ij}^{\beta}, k, l = 1, c \quad (6)$$

Supervised Learning

The goal of applying the classification algorithms is to determine the most appropriate data mining technique to predict the performance of the trainee teachers based on their data (Sen *et al*., 2020).

Decision Tree Classifier

The Decision Tree (DT) is a simple supervised learning algorithm implemented to solve regression and classification problems (Fletcher and Islam, 2020). The tree consists of a root, branches, and nodes. It is a model used to predict the class label of a sample by iteratively learning simple, easy-to-understand, and interpreting logical rules derived from the previous result. The decision tree leaves (branch tips) represent the class labels (decisions).

K-Nearest Neighbor Classifier

The principle of the k-nearest neighbor algorithm (Cunningham and Delany, 2022), often referred to as k-NN, is to predict the label of a new data point from a predefined constant number of k of training data closest in distance to that data point. In general, we use the Euclidean distance. Due to its ease of deployment and classification performance, the k-nearest Neighbor (k-NN) algorithm is commonly used in data mining.

Linear Discriminant Analysis

In this study, we will use another algorithm, named Fisher's Linear Discriminant Analysis (LDA) (Zhu *et al*., 2022), of well-known data mining used for dimensionality reduction and classification. The LDA classifier models the data conditional distribution for each class and uses Bayes' theorem. The model uses Gaussian distributions for each class with the same variance. Given data (an individual) to classify, we select the maximum posterior probability for each possible class.

Gaussian Naïve Bayes Classifier

Naive Bayes classifiers are a set of supervised learning algorithms that use Bayes' theorem. These naive probabilistic algorithms assume independence between features. In this study, we used the Gaussian naive Bayes Classifier (GNB) that uses maximum likelihood to estimate the mean and standard deviation (Rawal and Lal, 2023).

Support Vector Machine Classifier

For this classification problem, we applied a variant of the Support Vector Machine (SVM) algorithm which is a class of supervised learning methods based on maximizing the margin principle (separation of classes) (Pisner and Schnyer, 2020).

Random Forest Classifier

Random Forest (RF) is a method used to solve regression and classification problems (Parmar *et al*., 2019). It is a supervised learning algorithm; the forest is formed by mixing and combining several decision tree classifiers with improved predictive accuracy.

Data Collection Process

The given data set is a set of trainee teachers' information collected from two assessments in RCETP. It is difficult to evaluate all the individuals in the population (773 trainee teachers in the Fez-Meknes region) due to several constraints, such as time and financial constraints. In this study, we analyzed the results of evaluations of a sample consisting of trainee teachers belonging to the RCETP of Sefrou (179 individuals), we used two sampling techniques, for practical reasons of accessibility, the first is a convenience pilot sampling that consists of selecting a group of trainee teachers assigned to RCETP of sefrou (Edgar and Manz, 2017). The second type of sampling consists of dividing the sample resulting in the first stage into groups and then we randomly select a set of individuals from each group. It is worth noting that we reduced total and partial non-response by implementing a set of constraints. After this pre-processing, we were able to evaluate the responses of 135 trainee teachers, representing an exploitation rate of 75.41%.

The information-gathering process for the data set consists of two stages.

Gathering information from the diagnostic assessment distributed to 135 trainee teachers at the RCETP, fez-Meknes, sefrou, during the 2021-2022 season. We invited four experts to evaluate our diagnostic tool regarding content validity. According to Davis (1992); Waltz *et al*. (1991), the minimum number of experts was to be at least two assessors in the content area to be measured and at least one with knowledge of instrument construction. Other authors recommended several experts greater than three (Polit and Beck, 2006; Polit *et al*., 2007). First, we asked the experts to familiarize themselves with the concepts and skills that trainee teachers were expected to have as prerequisites, then they were asked to check whether the question was easily associated with the right skill, i.e., whether it represented it well. Then, the experts could make their judgment using the following scale:

- Not relevant
- Somewhat relevant
- Quite relevant
- Highly relevant

Finally, we calculated various content validity indices: Item CVI (I-CVI) and Scale level CVI (S-CVI). Two techniques are used to measure S-CVI: The average of the I-CVI scores (S-CVI/Ave) and the sum of universal agreement scores divided by the number of questions (S-CVI/UA). From the I-CVI we calculated the modified Kappa coefficient (Fleiss, 1981; Polit *et al*., 2007). The Fleiss kappa evaluation criteria are divided into four groups:

- Excellent (≥0.74)
- Good (0.60-0.73)
- Fair (0.40-0.59)
- Poor (≤0.39)

S-CVI/Ave and S-CVI/UA were 0.96 and 0.84 respectively, with a maximum I-CVI of 1.00 and a minimum of 0.75. The average modified kappa coefficient for the 40 questions (items) was 0.91, with a score of "excellent". An impressive 72.5% (n = 29) of items were rated excellent and 27.5% of questions were rated good (n = 11). These scores showed that there was acceptable agreement between these evaluators on the content validity of the 40 questions.

In this study, we also tested the purification of the measurement instruments (internal reliability of the data) using Cronbach's alpha.

The reliability of the information quality scale is 0.630 with 40 questions, a generally accepted rule is that an alpha of 0.6-0.7 indicates an acceptable level of reliability (Ursachi *et al*., 2015). We have eliminated some questions to increase reliability. Based on the results of 26 questions, with an alpha of 0.715, our scale demonstrates satisfactory reliability.

Gathering information from the scored activities of trainee teachers and the validation exam for the module (after 34 h of qualifying training). It should be noted that the summative assessment (the exam) is a problem situation proposed and validated by four domain experts.

In this study, the indices measured are acceptable and in addition, because of the type of concepts evaluated and the question's nature, we decided to use 26 questions instead of 40 for the diagnostic assessment and the exam proposed by the domain experts. Figure 1 shows the steps of the information collection process.

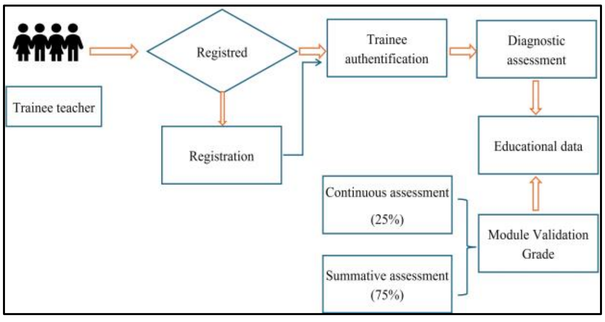

Fig. 1: Data collection process

Table 1: Main features in the dataset

Feature	Description	Type	Values before preprocessing
Sexe	The gender of the trainee	Nominal	Female or Male
Age	The age of the trainee	is strictly less than 30	
Home organization	The home institution of the trainee (faculty)	Nominal	"Juridical and chariaa sciences", "Economic Sciences", Letters and humanities", "Education", "Sciences"
DA Score	Diagnostic assessment score	Numeric	From 0-20
Module score	Module validation grade	Numeric	From 0-20

Table 2: Home institution and gender information (D: Juridical and chariaa sciences; E: Economic sciences; L: Letters and humanities; P: Education; S: Sciences)

	D	E	L	P	S	Total
Female	17	30	13	03	22	085
Male	15	14	05	04	12	050
Total	32	44	18	07	34	135
Percent (%)	23.70	32.59	13.34	5.18	25.19	100

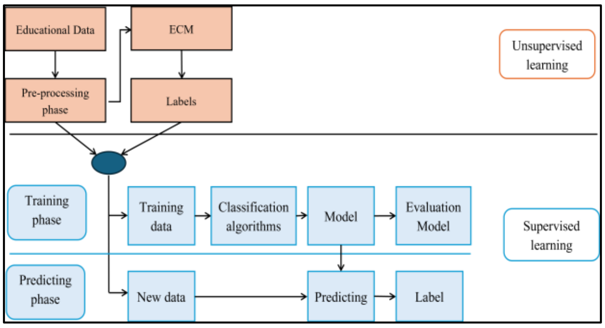

Fig. 2: Proposed model

Data (a trainee) consists of two features which are grades (diagnostic assessment grade, module validation grade) and other information about personal and academic characteristics (age, sex, home organization) as mentioned in Table 1.

The sample population is representative of the different disciplines in the various institutions of origin. It is noteworthy that there is an excessive representation of trainee teachers originating from management and economics institutions (32.59%), while trainee teachers from humanities institutions and institutions giving educational licenses represent only 13.33 and 5.18% respectively. Trainees from scientific and law institutions accounted for 25.18 and 23.70% respectively as presented in Table 2.

Data preparation is an essential part of data processing and it ensures the accuracy and relevance of the data to be used and therefore all data is pre-processed. It's important to note that we carried out further data pre-processing using the Factor Analysis of Mixed Data Method (FAMD), enabling the transformation of data containing both quantitative and qualitative variables (Pagès, 2004).

Firstly, we apply our ECM algorithm to the pre-processed data set with five features (age, sex, home organization, diagnostic assessment grade, and module validation grade). Secondly, we used the clusters obtained from the ECM algorithm as labels for the data and we split the data into train and test subsets, then 70% of the data was used as training data and 30% as test data. The main objective of supervised learning is to predict the label of a new input data (new trainee teacher data) during the predicting phase with high potential precision by starting with learning the labeled data in the training phase which classification algorithms are used to construct a suitable model that can precisely map the inputs to the desired outputs, as illustrated in Fig. 2.

In the model evaluation phase, a K-fold cross-validation technique was performed to evaluate and select the classification models. The K-fold cross-validation process consists of splitting the training dataset several times into k equally sized subsets (k folds) and the training-validation division is repeated k times (Marcot and Hanea, 2021). For each iteration, one fold is used as a validation set and the other k-1 folds are used as a training subset for model evaluation. We measure the Mean Square Error (MSE) of the validation process as the average of the k validation errors of each iteration defined below:

$$MSE = \frac{1}{k}\sum_{i=1}^{k}(L_i - M_i)^2 \quad (7)$$

where, L_i and M_i represent the i^{th} label and model prediction of the validation set respectively.

Table 3 shows different parameters and values used in classification algorithms.

DOI: 10.3844/jcssp.2024.1020.1029

Table 3: Algorithms parameters and values

Algorithms	Parameters	Values
DT	Criterion random_state max_depth	Gini none none
KNN classifier	n_neighbors weights	5 uniform
RF	n_estimators max_depth max_f features min_samples_split min_samples.Jeaf random_state	100 none sqrt 2 1 none
Gaussian NB	priors var_smoothing	None1e-9
Linear discriminant analysis	solver shrinkage priors n_components store_covariance tol covariance_estimator	svd none none none False 1.0e-4 none
SVC	C kernel degree gamma coef0 shrinking probability tol cache_size class_weight	1.0 rbf 3 scale 0.0 True False 1e-3 200 none

Results and Discussion

When using a clustering algorithm, it is a common practice to use the optimum number C of clusters. For this, the validation index N*(c) based on the non-specificity of a belief function, proposed in (Masson and Doneux, 2008), is used to select the C index of the ECM algorithm; this index must be minimal. ECM algorithm is tested with different cluster numbers with C = 2,3,4 and 5 and the parameters: $\delta = 20$ and $\beta = 2$ as shown in Fig. 3. We note that the optimal choice for the clustering outcome is equal to C = 2. Consequently, we implement our ECM algorithm with C = 2.

Figure 4 and Table 4 clearly show that the different clusters, obtained with the implementation of the ECM algorithm with C = 2, have different characteristics.

In this experiment, the ECM algorithm divided the trainee teachers into two groups representing singleton clusters. On the one hand, singleton cluster W1 (73 trainees) represents those who validated the module validation exam; 94.52% of trainee teachers scored well on the diagnostic assessment, while 5.48% scored close to 10 on the diagnostic assessment and excellent score on the module validation exam. We also note that 60.27% of trainees in cluster W1 came from management and economics institutions and 28.77% from scientific colleges. On the other hand, we also note that over 27.42% of trainees in the W2 singleton cluster (17 trainees) scored below 10 on the diagnostic assessment; most of these trainees came from juridical, chariaa, humanities, and letters colleges, we also note that all trainees in this cluster developed significantly and achieved an average score of 13.24 on the module validation exam. It should also be noted that trainees in cluster W1 evolved more than trainees in cluster W2.

These results clearly show that elementary school trainee teachers have different digital skills, indeed a significant percentage of them lack the necessary skills to create digital resources, share information, and communicate via learning and teaching platforms.

In our study, we examined trainee teachers' Information and Communication Technology (ICT) skills and their use of online platforms and resources. Three main hypotheses were formulated. The results presented above validate the hypotheses put forward in the introduction to this study. In our study, we assessed the hypothesis that trainees' ICT skills differ according to their home institutions (hypothesis 1). For this purpose, we analyzed trainees' scores on the diagnostic assessment, taking into account where the trainees came from. The results showed significant variations between the ICT skills of trainees from different establishments.

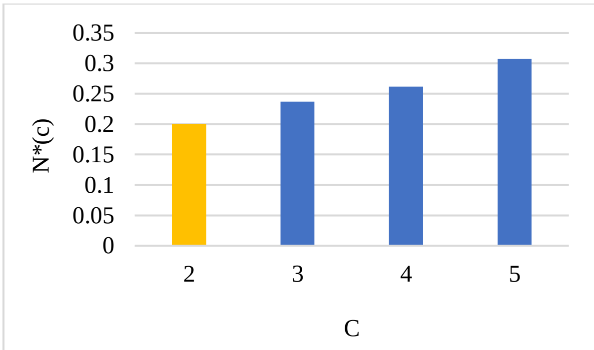

Fig. 3: Validation index N*(c) for variable cluster numbers with ECM algorithm

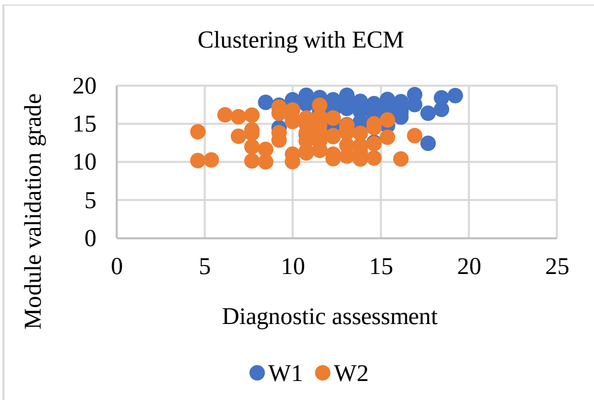

Fig. 4: Clustering results for trainee teachers using the ECM algorithm

Table 4: Clustering results

Cluster	N		Diagnostic assessment	Final exam
w_1	73	Mean	13.33	16.79
		Std	2.38	1.56
		Score >10	66.00	73.00
		Score = 10	3.00	0.00
		Score <10	4.00	0.00
w_2	62	Mean	11,09.00	13.24
		Std	2.84	2.33
		Score >10	39.00	61.00
		Score = 10	6.00	1.00
		Score <10	17.00	0.00

The outcomes allow us to validate the hypotheses put forward in the introduction to this study. Thanks to the sessions of the ICTE module, they were able to progress and improve their marks by validating the module.

We also note that 22% of trainees from juridical and Chariaa sciences institutions and 39% from letters and humanities institutions obtained lower scores (score below 10), in contrast to trainee teachers from scientific, economic, and pedagogical institutions achieving good results and attaining excellence in the development of digital skills in the educational context. Indeed, only 9% of trainees from scientific and economic institutions obtained a score below 10, while all trainees from institutions delivering pedagogical licenses obtained a score above 10. In fact, despite the low initial scores obtained by some trainees during the diagnostic assessment, these difficulties were gradually overcome thanks to the curriculum adopted.

After validating hypothesis 1, according to which the ICT skills of trainee teachers vary according to their home institutions, it is crucial to consider concrete actions to reinforce these skills within university programs. The results thus underline the importance of rethinking these programs to effectively integrate technologies. This calls for an overhaul of academic programs to include a significant component dedicated to the development of information and communication technology skills.

The results also show that the trainees had ICT skills or had used or rarely used with their teachers (hypotheses 2 and 3) the platforms and resources during the COVID-19 pandemic.

These three hypotheses were largely validated in our study since the clustering algorithm revealed that a large majority of trainee teachers (84.44%), who responded to the diagnostic test, had ICT skills. While 15.56% lacked ICT skills and scored low.

It is important to note that all trainee teachers passed the module validation exam after 34 h of training at the RCETP. The results showed that there was a difference between the trainee teachers' achievements on the diagnostic test (some trainees lacked knowledge of information and communication technologies) and the post-test (module validation exam), which means that training at the center had a positive effect on the trainees' level. In this study, we used the K-fold cross-validation process (K = 10) to select the optimal model and classifier (Ghio *et al*., 2012; Nti *et al*., 2021).

The graph in Fig. 5 provides a visual overview of the evolution of the mean square error as a function of the classifier. This provides an excellent visualization of the relative performance of the different classification algorithms. By observing how the mean square error varies for each classifier, we can identify general trends and significant variations that might indicate differences in performance between the algorithms. Table 5 completes this analysis by providing a detailed comparison of the classification results obtained using the different classifiers. This step-by-step comparison enables an in-depth assessment of the robustness of each algorithm on our dataset. By making use of both plot and tabular analysis, we can reach informed decisions on which classifier is most appropriate for our application.

According to the result depicted in Fig. 5 and Table 5, the optimal classifier that gives the minimum error is the decision tree classifier. From the obtained results, it's clear that the decision tree classifier beats the other algorithms in terms of accuracy and error, closely followed by the other algorithms.

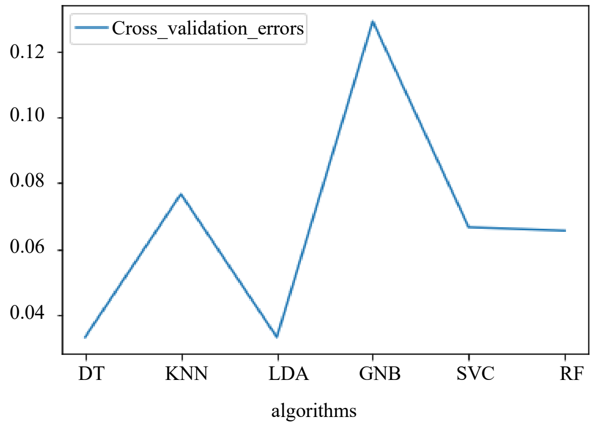

Fig. 5: Cross-validation error for different algorithms

Table 5: Performance of different classification algorithms for trainee teacher's dataset

Algorithms	Training accuracy	Testing accuracy
DT	1.0000	0.9512
k-NN	0.9574	0.9268
LDA	0.9681	0.9756
GNB	0.8936	0.9512
SVC	0.9681	0.9512
RF	1.0000	0.9512

DOI: 10.3844/jcssp.2024.1020.1029

Figure 6 shows the confusion matrix for the selected classifier (decision tree classifier), which provides a visual representation of the classifier's performance in terms of differentiating between the various classes. Each grid cell depicts the number of individuals predicted to belong to a certain class, compared with the actual class to which they belong. This is essential for understanding the behavior of the classifier, as it shows the levels of confusion between the different classes.

An analysis of the confusion matrix's diagonal values indicates the number of correctly predicted cases for each class. In this context, 39 out of a total of 41 trainee teachers were correctly classified, indicating the model's high performance. As a result, the overall accuracy of the classifier rises to an impressive 95.12%, testifying to its ability to predict correctly.

However, accuracy alone may not provide a complete assessment of model performance. It is therefore essential to take into account other measures such as precision, recall, and F1-score. The macro-mean, or macro-average, of these measures, gives a more detailed view of the model's effectiveness. In this model, the arithmetic mean (macro avg) between the F1-score of the 2 classes is equal to 0.95. That means the model performs in a balanced way in terms of precision and recall for the different classes. A high F1 score indicates that the model is capable of making accurate predictions at the same time by reducing false positives and false negatives. In addition, Table 6 presents a detailed summary of the various measures used to evaluate the model's performance. These measures provide an overview of the model's ability to generalize and make accurate predictions on novel data. By carefully analyzing these metrics, all parties concerned can make informed decisions about deploying and optimizing the classifier in other real-life situations. The combination of accuracy, precision, recall, and F1-score gives a comprehensive assessment of the classifier's performance and appropriateness for a particular task.

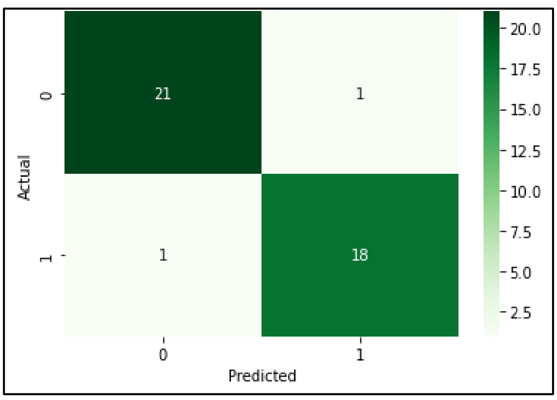

Fig. 6: Confusion matrix without normalization for decision tree classifier

Table 6: The classification reports

	Precision	Recall	F1-score	Support
C1	0.95	0.95	0.95	22
C2	0.95	0.95	0.95	19
Accuracy			0.95	41
Macro avg	0.95	0.95	0.95	41
Weighted avg	0.95	0.95	0.95	41

Conclusion

The aim of this study was to build a hybrid model that implements the evidential C-Means algorithm in the first place and classification algorithms in the second place in order to group trainee teachers with similar statistical characteristics and predict their performance. Although this study focused on a single class of 2021-2022 and a larger sample size is essential, the results obtained are relevant and remarkable. The analysis showed that trainees from humanities institutions do not possess sufficient ICT skills and perform poorly on the diagnostic assessment. We also find that trainees from rights institutions score lower, in contrast to trainee teachers from scientific, economic, and educational institutions, who score high and achieve excellence in the development of digital skills in the educational context. There are several reasons for these results, the first and most significant being linked to the trainees' educational background. The second reason concerns the factors that influence the integration of ICT into higher education disciplines, such as contextual, social, pedagogical, personal, and institutional characteristics. The ICTE module sessions enabled them to make progress and they improved their marks by validating the module. In fact, despite initial low scores on the diagnostic assessment, these difficulties were gradually overcome through the curriculum adopted at the training center. This study also aimed to implement and compare different classification algorithms in terms of accuracy rates by building a model to predict the performance and results of trainee teachers. The integration of clustering and classification algorithms into our hybrid model opens up a new realm in educational data analysis, offering transformative opportunities and actionable strategies for improving learning performance. Evaluation of the implemented classification algorithms, including linear discriminant analysis, random forest, Support Vector Machines (SVM), GN, DT, k-NN classifier, and Gaussian NB, highlighted the importance of algorithm choice in hybrid model performance. Each algorithm offers its own unique advantages and benefits. However, the selection of the algorithm should be based on the characteristics of the dataset. We will envisage the optimization of this model and the use of other data mining approaches that model the thinking process of trainee teachers and predict their performance.

Acknowledgment

We sincerely thank the editor for his immense support during the publication process of this research article. In addition, we are grateful to the editorial team for their diligent efforts in reviewing and editing our work, and we are gratified by the opportunity to make a significant contribution to the field of research via this publication.

Funding Information

The authors have not received any financial support or funding to report.

Author's Contributions

Yissam Lakhdar: Conceptualization, methodology, software, validation, formal analysis, written-original draft prepared, written-review and edited, visualized. Resources, data curation, investigation, supervision and project administration.

Khawla El Bendadi and Brahim Bakkas: Conceptualized, methodology, software, validation, formal analysis, written-original drafted prepared, written-review and edited, visualized.

Ethics

This article is original and unpublished. Correspondence authors confirm that all other authors have read and agree that the manuscript does not involve ethical issues.

References

Albreiki, B., Zaki, N., & Alashwal, H. (2021). A Systematic Literature Review of Student' Performance Prediction Using Machine Learning Techniques. *Education Sciences*, *11*(9), 552. https://doi.org/10.3390/educsci11090552

Bachiri, H., & Sahli, R. (2020). The Need for Distance Learning in the Wake of COVID-19 in Morocco: *International Journal of Language and Literary Studies*, *2*(3), 240–256. https://doi.org/10.36892/ijlls.v2i3.326

Berrada, K., El Kharki, K., & Ait Si Ahmad, H. (2022). Science Education in Morocco. In R. Huang, B. Xin, A. Tlili, F. Yang, X. Zhang, L. Zhu, & M. Jemni (Eds.), *Science Education in Countries Along the Belt & Road* (1st Ed., pp. 93–112). Springer Nature Singapore. https://doi.org/10.1007/978-981-16-6955-2_7

Bezdek, J. C., Ehrlich, R., & Full, W. (1984). FCM: The fuzzy c-means clustering algorithm. *Computers & Geosciences*, *10*(2–3), 191–203. https://doi.org/10.1016/0098-3004(84)90020-7

Chen, J., Zhou, G., & Zhong, Y. (2024). A personalized recommendation system for teaching resources in sports using the fuzzy C-means clustering technique. *Soft Computing*, *28*(1), 703–720. https://doi.org/10.1007/s00500-023-09421-0

Cunningham, P., & Delany, S. J. (2022). k-Nearest Neighbour Classifiers - A Tutorial. *ACM Computing Surveys (CSUR)*, *54*(6), 1–25. https://doi.org/10.1145/3459665

Dave, R. N. (1991). Characterization and detection of noise in clustering. *Pattern Recognition Letters*, *12*(11), 657–664. https://doi.org/10.1016/0167-8655(91)90002-4

Davis, L. L. (1992). Instrument review: Getting the most from a panel of experts. *Applied Nursing Research*, *5*(4), 194–197. https://doi.org/10.1016/s0897-1897(05)80008-4

Edgar, T. W., & Manz, D. O. (2017). *Research methods for cyber security*. Syngress. ISBN: 9780128053492

Fleiss, J. L. (1981). *Statistical Methods for Rates and Proportions*. Wiley, John and Sons, Incorporated. https://doi.org/10.1002/0471445428

Fletcher, S., & Islam, Md. Z. (2020). Decision Tree Classification with Differential Privacy. *ACM Computing Surveys (CSUR)*, *52*(4), 1–33. https://doi.org/10.1145/3337064

Ghio, A., Anguita, D., Oneto, L., Ridella, S., & Schatten, C. (2012). Nested Sequential Minimal Optimization for Support Vector Machines. *Artificial Neural Networks and Machine Learning–ICANN 2012: 22nd International Conference on Artificial Neural Networks*, 156–163. https://doi.org/10.1007/978-3-642-33266-1_20

HCETS. (2015). *For a School of Equity, Quality and Promotion. A Strategic Vision of Reform 2015-2030*.

Hoot, J. L., Szente, J., & Mebratu, B. (2004). Early Education in Ethiopia: Progress and Prospects. *Early Childhood Education Journal*, *32*(1), 3–8. https://doi.org/10.1023/b:ecej.0000039637.29327.d2

Krishnapuram, R., & Keller, J. M. (1996). The possibilistic C-means algorithm: insights and recommendations. *IEEE Transactions on Fuzzy Systems*, *4*(3), 385–393. https://doi.org/10.1109/91.531779

Lamichhane, C. D. (2018). Understanding the Education Philosophy and Its Implications. *NCC Journal*, *3*(1), 24–29. https://doi.org/10.3126/nccj.v3i1.20245

Marcot, B. G., & Hanea, A. M. (2021). What is an optimal value of k in k-fold cross-validation in discrete Bayesian network analysis? *Computational Statistics*, *36*(3), 2009–2031. https://doi.org/10.1007/s00180-020-00999-9

Masson, M.-H., & Denœux, T. (2008). ECM: An evidential version of the fuzzy c-means algorithm. *Pattern Recognition*, *41*(4), 1384–1397. https://doi.org/10.1016/j.patcog.2007.08.014

Mounjid, B., El Hilali, E., Amrani, F., & Moubtassime, M. (2021). Teachers' Perceptions and the Challenges of Online Teaching/Learning in Morocco during Covid-19 Crisis. *Arab World English Journal*, *7*(1), 38–52. https://doi.org/10.24093/awej/call7.3

Nafea, I. T. (2018). *Machine Learning in Educational Technology* (H. Farhadi, Ed.; 1st Ed.). InTech. https://doi.org/10.5772/intechopen.72906

Nejjari, A., & Bakkali, I. (2017). L'usage des TIC à l'école marocaine: état des lieux et perspectives. *Hermès*, *78*(2), 55. https://doi.org/10.3917/herm.078.0055

Nti, I. K., Nyarko-Boateng, O., & Aning, J. (2021). Performance of Machine Learning Algorithms with Different K Values in K-fold CrossValidation. *International Journal of Information Technology and Computer Science*, *13*(6), 61–71. https://doi.org/10.5815/ijitcs.2021.06.05

Outoukarte, I., Ben Fares, S., Itouni, H., Rassou, K. K., & Tahiri, A. (2023). Distance learning in the wake of COVID-19 in Morocco. *Heliyon*, *9*(6), e16523. https://doi.org/10.1016/j.heliyon.2023.e16523

Pagès, J. (2004). Analyse Factorielle de Données Mixtes: Prinbcipes et Exemple d'application. *Revue de Statistique Appliquée*, *52*(4), 93–111. https://eudml.org/doc/106558

Parmar, A., Katariya, R., & Patel, V. (2019). A Review on Random Forest: An Ensemble Classifier. *International Conference on Intelligent Data Communication Technologies and Internet of Things (ICICI) 2018*, 758–763. https://doi.org/10.1007/978-3-030-03146-6_86

Pisner, D. A., & Schnyer, D. M. (2020). Support vector machine. In *Machine Learning* (1st Ed., pp. 101–121). Elsevier. https://doi.org/10.1016/b978-0-12-815739-8.00006-7

Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? critique and recommendations. *Research in Nursing & Health*, *29*(5), 489–497. https://doi.org/10.1002/nur.20147

Polit, D. F., Beck, C. T., & Owen, S. V. (2007). Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Research in Nursing & Health*, *30*(4), 459–467. https://doi.org/10.1002/nur.20199

Raouf, K., Boulahoual, A., Mezzat, F. Z., & Sahli, M. (2020). Les TICE au Maroc Entre instauration technologique et conduite de changement, cas des établissements scolaires. *European Scientific Journal ESJ*, *16*(4), 89–106. https://doi.org/10.19044/esj.2020.v16n4p89

Rawal, A., & Lal, B. (2023). Predictive model for admission uncertainty in high education using Naïve Bayes classifier. *Journal of Indian Business Research*, *15*(2), 262–277. https://doi.org/10.1108/jibr-08-2022-0209

Sen, P. C., Hajra, M., & Ghosh, M. (2020). *Supervised Classification Algorithms in Machine Learning: A Survey and Review* (pp. 99–111). Springer Singapore. https://doi.org/10.1007/978-981-13-7403-6_11

Sen, S., & Dave, R. N. (2002). Clustering of relational data containing noise and outliers. *1998 IEEE International Conference on Fuzzy Systems Proceedings. IEEE World Congress on Computational Intelligence (Cat. No.98CH36228)*, 1411–1416. https://doi.org/10.1109/fuzzy.1998.686326

Ursachi, G., Horodnic, I. A., & Zait, A. (2015). How Reliable are Measurement Scales? External Factors with Indirect Influence on Reliability Estimators. *Procedia Economics and Finance*, *20*, 679–686. https://doi.org/10.1016/s2212-5671(15)00123-9

Waltz, C. F., Strickland, O., & Lenz, E. R. (1991). *Measurement in Nursing Research*. FA Davis Company. https://lccn.loc.gov/2016015495

Zhou, Z. H. (2021). *Machine Learning*. Machine Learning. Springer Singapore. https://doi.org/10.1007/978-981-15-1967-3

Zhu, F., Gao, J., Yang, J., & Ye, N. (2022). Neighborhood linear discriminant analysis. *Pattern Recognition*, *123*, 108422. https://doi.org/10.1016/j.patcog.2021.108422